



## Face Recognition based on Hidden Markov Model and Canny Operators

Sameh Magdy, Tarek Mahmoud, Mohamed Ibrahim  
*School of Communication Engineering, Canadian international college*  
 Sameh\_mohamed@cic-cairo.com,  
<https://www.cic-cairo.com/cic/>

### Abstract

In this paper, a new face recognition technique based on Hidden Markov Model (HMM), Pre-processing, and feature extraction (Canny) is proposed. Two main contributions are presented. In the pre-processing, the input image's edge is fixed to five pixels and spacing between various parts of the pixels is fixed to a trial threshold. The second contribution is a new technique to extract the image's features by splitting the image into non-uniform height depending on the distribution of the foreground pixels. The foreground pixels are extracting by using the vertical sliding windows. The pre-processing steps have enhanced the performance of the system, for example, speed is the essence. In this paper, Hidden Markov models alongside the canny operator which proved high accuracy recognition of the face and its boundary. The presented technique is faster than some other techniques that are investigated for comparison. In addition to that, it reveals, its ability to recognize the normal face and face boundary very efficiently

**Keywords:** HMM, Canny operator, Face Recognition, Accuracy

### Nomenclature

DCT	Discrete Cosine Transform
HMM	Hidden Markov Model
LDA	Linear discriminate Analysis
ORL	Olivetti Research Laboratory
PCA	Principal Component Analysis
PDBNN	Probabilistic Decision- Based Neural Network
SVD	Singular Values Decomposition
$W_o$	represents the width of overlapping of the vertical frames

### 1. Introduction

Many up to date techniques have been immersed during the recent years, due to the urgent need for security issues and in particular the pattern recognition. In digital image processing, edge detection [1] is considered the main element of recognition method. Therefore, it is a necessary for image analysis to extract its landmarks, forms, and then convert it into data that can be processed. Face recognition has become necessary in commercial,

law and security researches. Recently, a dedicated work targets face recognition process that aims to handle the relation between time and accuracy. However, face recognition faces too many challenges such as the difference between human faces in color, shape, and dimensions. Thus, a demand for image analysis has appeared to extract its landmarks and forms, and then convert it into data to be processed. A different number of models can be used for addressing this issue. Such as, stochastic process is the hidden Markov model. There are two main features used for face detection. The first one is the feature based where analysis is performed on the main geometry of the face like eyes, nose, ears, and mouth. The second feature is a holistic base where the face is analyzed as a two-dimensional pattern. The available researches have followed a certain techniques for data reduction and feature extraction, which is used for face detection. A wide range of models can be used for addressing this issue, most commonly models are Geometrical Feature Matching [2], Eigenfaces [3], Neural Networks [4], Linear discriminate analysis (LDA) [5], Principal component analysis (PCA), Discrete cosine transform (DCT) [6] and the fuzzy logic technique [7], which is used for face detection and Hidden Markov Models. Geometrical feature matching model is based on extracting geometrical features to form a complete picture of the face. This model showed a low accuracy with 75% on ORL database recognition rate. Another model is Eigenfaces method which uses the Principal Component Analysis (PCA) with 90.5% recognition rate on ORL database. Although this model is simple and fast however it is not robust enough over the changes in face orientation and lighting conditions. Hidden Markov Model is used for developing features extraction methods and enhancing not only the operation of face detection but also the data processing time as will be shown in this paper. In this paper we introduce a new approach for preprocessing the input image and a new method of extracting features by splitting the input image vertically with non-uniform heights. This paper is sorted as follows. In section 2, the proposed (presented) technique is explained. In section 2.1, the pre-processing step is presented. In section A, the feature extraction followed by recognition system with the use of HMM is



introduced in section B and C. In section 3 simulation results are described. In section 4 discussion is presented. Finally, conclusions are highlighted in section 5

## 2. Material and Methods

Material used in our proposed recognition technique:

- Available subsets of the benchmark ORL face database. In our technique validation, we used 80 images for training purpose (8 persons  $\times$  10 images), and 40 images for testing purpose (8 persons  $\times$  5 images). The face recognition technique presented in this paper was developed, trained, and tested.
- Using Matlab 7.10, the computer was a Windows 7 machine with a 2.1 GHz Intel core 2 processor and 2.87 GB of RAM.

### 2.1. Methodology

Our proposed Technique consists of multiple steps:

- The pre-processing step which will be explained in this section.
- The feature extraction to prepare data to be processed.
- Applying Canny operator
- Applying HMM on the recognition system.
- We calculate the accuracy (recognition rate) to measure the ability of recognize the image by the system. Figure 1 shows the block diagram for our proposed technique.

#### Pseudo code

```

READ new input image
not_known = 1
WHILE (not_known)
    DO median filter to reduce noise
    DO other enhancement
    DO Canny to detect edges
IF (image_size! = threshold size)
    DO resize the image to new size
ENDIF
DO segmentation of the resized image with
overlapping sliding window (horizontal and
vertical)
DO extract features from the segments of the image
IF (HMM_training)
IF (new_class)
    DO add new class model
    ELSE
DO modify HMM class model using the image
ENDIF
ELSE
Classify the image using the HMM models
IF recognized
    WRITE output the image with decision
    Break from while loop
ENDIF
ENDIF
ENDWHILE

```

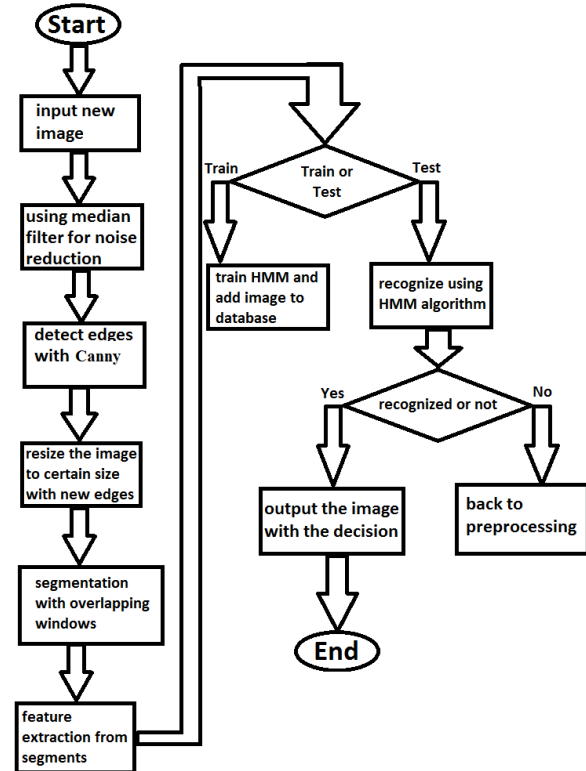


Figure 1 Face Recognition Proposed Technique

### A. Pre-processing

The image from RGB format is change to analogous grayscale format, as a first step. A grayscale image is a monochrome image, which contain data related to brightness only without containing any information about the color. A grayscale data represents a broad range of intensities. The typical image contains 8 bit/pixel, this lets the image to be represented with different level of brightness (0-255) [8]. The pre-processing is mandatory step to enhance the ease of features extraction and reduce the noise in the image, by using the median filter. The pre-processing step includes changing of the image's edges to a settled pixel's numbers, that occasionally decrease the space between parts of pixels, if this space is larger than a calculated threshold. In pre-processing we change the input image's edges, There are contradictions in the image's edges in the database. After several trials, it was found that a rise in the recognition rate after normalizing the input image's edges and pixels, that goes through two steps:

- Change the input image's edges as shown in Figure 2.b; after changing, the image's edges in the original are thinner than others of the same image as shown in Figure 2.a. After thinning, the image's edges of each will have become the same pixels (2 pixels) as shown in Figure 2.c.
- Spread the binary image by using 4-pixels framing the original pixel, as shown in Fig. 2.d. So, the two image's edges will become the same.



In pre-processing, we decrease the space between pixel's parts to find the places where no pixels exist so the vertical projection is used, after that calculates the width (d) of each place, as shown in figure 3a. If the width is larger than a threshold (8 pixels, by trial and error), then removing (d-threshold) to get a stable separation between of pixel's parts as shown in Figure 3.b

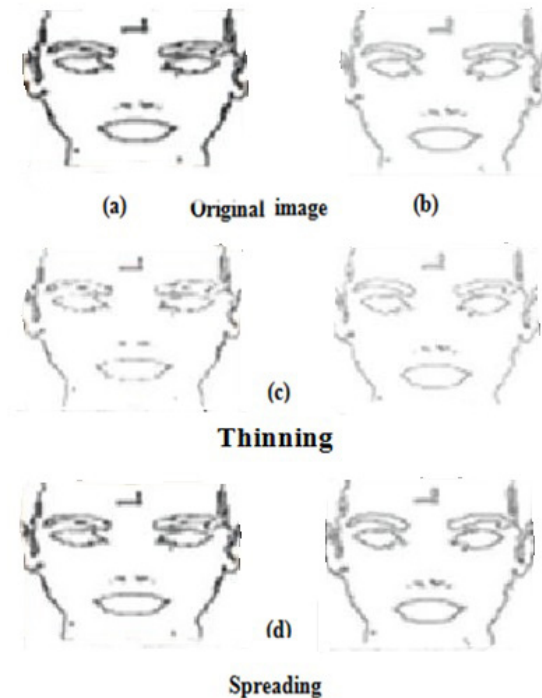


Figure.2 The input image's edges are changing, in (a, b) the identical original image, which the edges are not the same, (c) After thinning the edges, (d) After spreading to fix the image's edges.

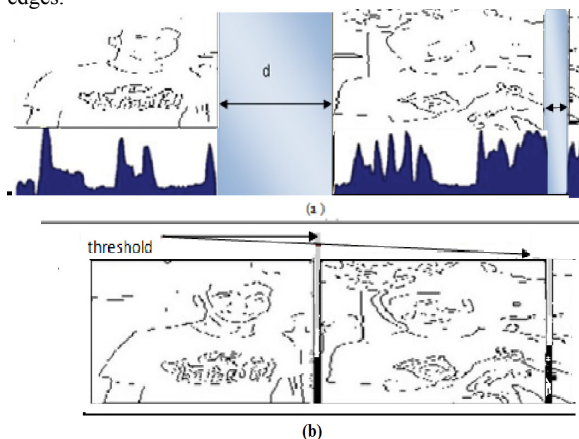


Figure 3 (a) Original image with large space between pixel's parts, (b) Decreasing the space between pixel's parts.

## B. Feature Extraction

After the pre-processing stage, the binary format of the input image as represented in the ORL database, which is used to get a feature vector. Identified the pixel's background "zero" and the foreground "one". The image is split into  $n$  horizontal segments, which have an identical number of the pixel located in the foreground in

each segment. The total pixels of the image are counted and divided by  $n$ , to get the exact number of pixels per segment. The image is split into vertical frames (overlapping), each frame is split into fixed-width small cells, that the cell's height reliant on the distribution and the pixels of the foreground in the images. In the image, the sliding window is moved from right to left and determines the features for each frame, as shown in Figure 4

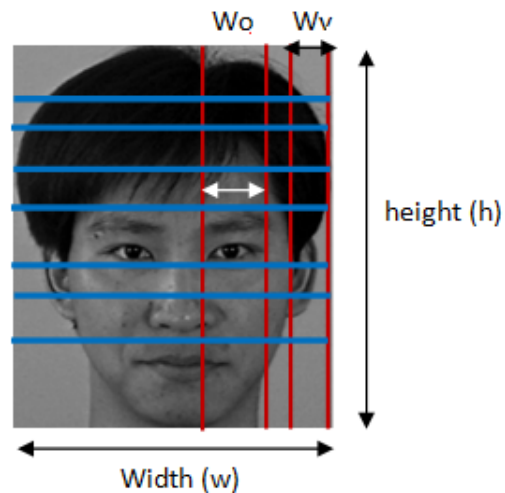


Figure.4 Using the sliding window

The image's height represents ( $h$ ), the image's width represents ( $W$ ), ( $W_o$ ) represents the width of overlapping of the vertical frames, and ( $W_v$ ) represents the width of the vertical frame. The best scenario in our technique when ' $W_v = 7$  pixels with ' $W_o = 4$  pixels, this sliding accomplished the 100% accuracy (recognition rate) as it extracted the maximum necessary features for recognition. As it extracted the maximum necessary features for recognition. 6 cells (trial and error) are used per vertical frame with non uniform heights to extract this feature for each cell, the number of "1" pixels are calculated to produce 6 features per frame (6 cells per frame x 1 feature per cell) as shown in figure5.



Figure5 Six features per frame by dividing the image into six non uniform heights



## • Canny Edge Detection

### Advantage of Canny:

- Good detection accuracy
- Good localization:
- Single response constraint

Canny edge detection can be applied using the following five steps [9]:

- 1) Removing the noise and smooth the images by applying Gaussian filter.
- 2) Finding the intensity gradients of the image.
- 3) Getting rid of spurious response to edge detection by applying non-maximum suppression.
- 4) Determine potential edges by applying double threshold.
- 5) Track edge by hysteresis: Remove all weak edges that are not connected to strong edges to finalize the process of detection.

Detector eliminates noise by smoothing the image, and then it finds the gradient to show high spatial derivatives regions [6]. In these areas the canny edge detector performs tracking and eliminates any pixel that is not at the maximum. [7].

### Algorithm

1- Compute  $f_x$  and  $f_y$

$$f_x = \frac{\delta}{\delta x} (f * G) = f * \frac{\delta}{\delta x} G = f * G_x \quad (1)$$

$$f_y = \frac{\delta}{\delta y} (f * G) = f * \frac{\delta}{\delta y} G = f * G_y$$

$G(x, y)$  is the Gaussian function,  $G_x(x, y)$  is the derivate of  $G(x, y)$  with respect to  $x$  :

$$G_x(x, y) = \frac{-x}{\sigma^2} G(x, y) \quad (2)$$

$$G_y(x, y) = \frac{-y}{\sigma^2} G(x, y) \quad (3)$$

2-Compute the gradient magnitude:

$$magn(i, j) = \sqrt{f_x^2 + f_y^2} \quad (4)$$

3- Apply non-maxima suppression and hysteresis thresholding/edge linking.

## C. Hidden Markov Model

HMM process is a complex stochastic process. This process is repeated twice to form a well- defined model for the sequential data. HMM is a defined set of states, and transition probabilities control the transitions among those states. HMM consists of a different number of states and transition between those states. These transitions are controlled by transitions probability distribution. HMM process considers the feature observations in every state, those to produce the confidence of the modeled object according to its probability density function. From the unknown pattern, the classification can be obtained according to the extracted feature by searching a model that produces a maximum probability [10]. In this paper, HMM is discrete. In HMM Viterbi algorithm is searching for the best sequence of a hidden state given the input feature vector.

Three states HMM are represented in fig.5, which shows that we let a move to the current and the following states only. Two of these are emitting states and have their output probability distributions [11]. The transition matrix for this model is  $[3 \times 3]$ , which represent the transitions between these states such that in the case of transition, we will set the matrix element to one and matrix element will be set to zero in the case of no transition.

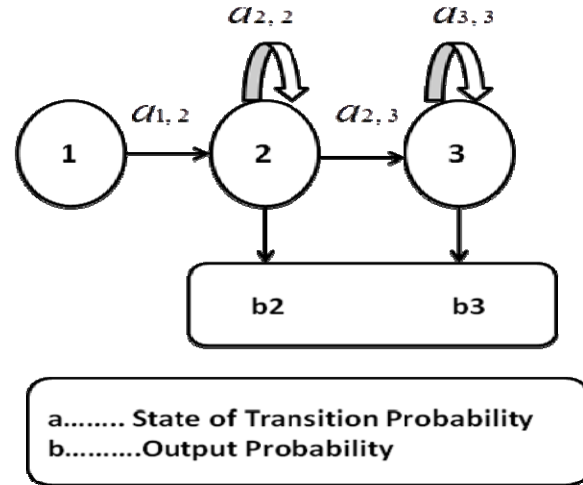


Figure 6 Three states of transition from left to right

The element of HMM are:

- $N$  is the number of state in the model. If  $S$  is the set of state, then  $S = \{S_1, S_2, S_3, \dots, S_N\}$ . The state of the model at time  $t$  is given by  $q_t \in S, 1 \leq t \leq T$ , where  $T$  is the length of the observation sequence (number of frames).
- $M$ , the number of different observation symbols. If  $V$  is the set of all possible observation symbols  $V = \{v_1, v_2, \dots, v_M\}$ .
- A the state of transition probability matrix i.e.  $A = \{a_{ij}\}$  where

$$a_{ij} = P[q_t = S_j \mid q_{t-1} = S_i] \quad (5)$$

$$1 \leq i, j \leq N \quad \text{with the constraint, } 0 \leq a_{i,j} \leq 1$$

and,

$$\sum_{j=1}^N a_{ij} = 1, \quad 1 \leq i \leq N$$

- $B$ , the observation of symbol probability matrix i.e.  $B = \{b_j(k)\}$ , where

$$b_j(k) = P[O_t = v_k \mid q_t = S_j] \quad (6)$$

$1 \leq j \leq N, 1 \leq k \leq M$  and  $O_t$  is the observation symbol at time  $t$ .

- $\pi$ , the initial state distribution, i.e.  $\pi = \{\pi_i\}$  where  $\pi_i = P[q_1 = S_i], 1 \leq i \leq N$  (7)

Using a notation, a HMM is defined as the triplet





$$\lambda=(A,B,\pi) \quad (8)$$

the above characterization corresponds to a discrete HMM. In a continuous density HMM, the state are characterization by continuous observation density functions. The most general representation of the model probability density function (PDF) is a finite mixture of the form:

$$b_i(o)=\sum_{k=1}^M C_{ik}N(O, \mu_{ik}, U_{ik}) \quad (9)$$

where,  $C_{ik}$  is the mixture coefficient for the Kth mixture in state i. without losses of generality  $N(O, \mu_{ik}, U_{ik})$  is assumed to be a Gaussian PDF with mean vector  $\mu_{ik}$  and covariance  $U_{ik}$

### 3. Simulation Results

The following parameter has been determined through validation: Number of states. Figure6 shows the samples of ORL face database. The following table 1 the states and effect of changing time versus accuracy (recognition rate). Table 2 and 3 shows the comparison with different techniques in recognition rate and Training Time. Figure7 shows the sample of ORL database Figure8, 9 shows the effect of the noise and PSNR. Figure 10 shows the effect of the given data in table1, Figure.11 shows the proposed technique is proved to be fast with respect to some other techniques that are investigated for comparison. Table 4 the proposed method which tested on images size of ORL database where the other methods use the same images size and measuring the error percentage after testing process.



Figure 7 Sample of ORL face database

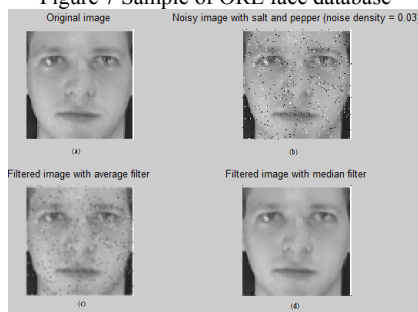


Figure 8 The effect of noise in sample ORL face database  
a-Original Image b- Gaussian noise image  
c- Average filter d- median filter

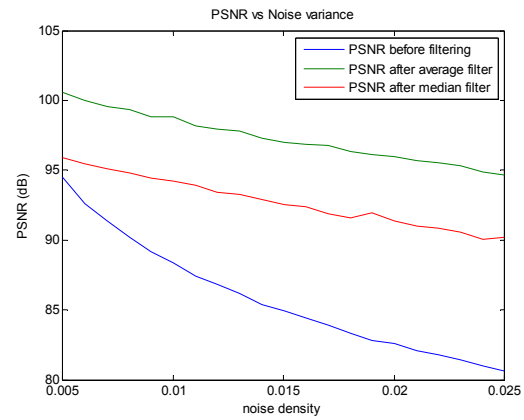


Figure 9 PSNR before and after different filters.

Table 1: shows the effect of timing and states on the recognition rate (accuracy) results

#sample	Average Time/Image(sec)	Recognition rate(Accuracy %)
1	0.00331	80.4
2	0.00336	83.3
3	0.00339	85.1
4	0.00346	86.6
5	0.00351	88.3
6	0.00360	92.3
7	0.00366	94.2
8	0.00370	98.2
9	0.00372	98.5
10	0.00376	99.2
11	0.00380	100
12	0.00381	100
13	0.00381	100

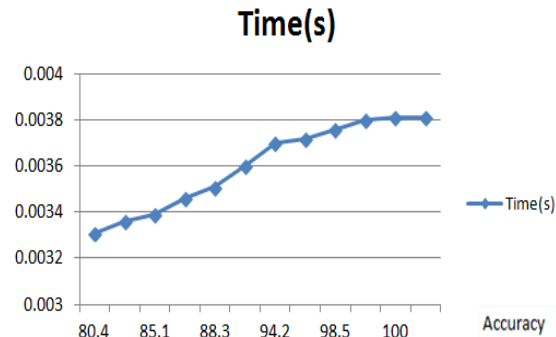


Figure. 10 Times (ms) versus the Accuracy

Table2 Comparison Results of Different Face recognition algorithms

Models	Recognition rate (%)
<b>Proposed Technique</b>	<b>100</b>
E-HMM	99.16
Eigenfaces	85.9
Fisherfaces	92.3
Laplacianfaces	93.2
ML-HMM	88.2
MCE-HMM	93.3
MC-HMM	97.5



Our proposed technique gives the highest recognition accuracy for all the sample sets in training and testing images. Comparison result are shown in figure11 and Table3.

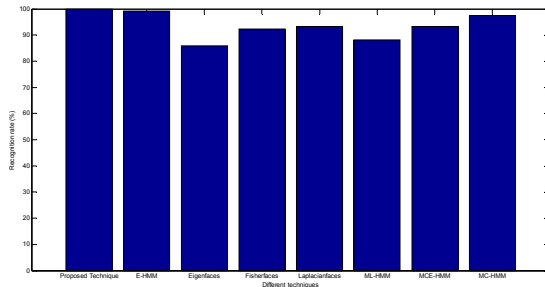


Figure11 Comparison between different techniques

Table 3 Comparison ORL database with different technique in Recognition rate and training time

Model	Recognition Rate (%)	Training Time/ image
<b>Proposed Technique</b>	<b>100%</b>	<b>0.00300 sec</b>
PDBNN	96%	20 min
n-tuple	86%	0.9
1DHMM + Wavelet	100%	1.13 sec
Pseudo-2D HMM	95%	1 sec
DCT-HMM	99.5 %	23.5 sec
1D HMM + SVD	99 %	0.63 sec

The Comparison with different face recognition techniques and the proposed technique on the ORL face database. It is important to notice that all different face recognition techniques in Table 3 use 112×92 resolution of ORL face database where we use 64×64 image size.

Table 4 Comparative results on ORL database, which is tested on images size of ORL database where the other methods use the same images size.

Method	Error
<b>Proposed Technique</b>	<b>0%</b>
Top-down HMM + gray tone features	13 %
Eigenface	9.5 %
Pseudo 2D HMM + gray tone features	5%
Elastic matching	20%
PDNN	4%
Continuous n-tuple classifier	2.7%
Top-down HMM + DCT coef.	16%
Point-matching and correlation	16%
Ergodic HMM + DCT coef.	0.5%
Pseudo 2D HMM + DCT coef.	0%
SVM + PCA coef.	3%
Independent Component Analysis	15%
Gabor filters + rank correlation	8.5%
Pseudo 2D HMM + Wavelet	0%
Markov Random Fields	13%
1D HMM+SVD	1%

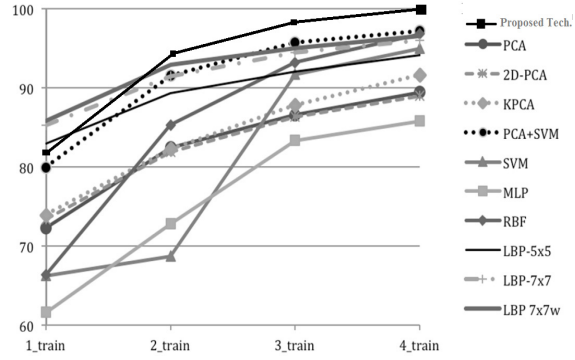


Figure12 Comparison between different techniques

It is worth mentioning the technique has been implemented without performing the pre-processing step mentioned in section A. The results for different sets of databases showed that the recognition rate is about 97%. This implies clearly the importance of the pre-processing on the accuracy (recognition rate).

#### 4. Discussion

The recognition rate of the proposed technique in this paper is higher than that of the traditional algorithms based on HMM. Also, we compare our proposed methodology with different face recognition methods such as EigenFace, FisherFace, Laplacianfaces, ML-HMM, MCE-HMM, and MC-HMM [12], our proposed method gives the highest recognition accuracy for all the sample sets (training and testing). Comparison result is summarized in the Table2, 3 and 4 and Figure 11.

It is worth mentioning the technique has been implemented without performing the pre-processing step mentioned before. The results for four different sets of databases showed that the recognition rate is about 97%. (I.e. Images in ORL database passed through several stages:

- 1- Improve the signal to noise ratio by using the filters (average and median filter).The PSNR before pre-processing at different faces shown in figure8
- 2- Normalizing the input image's edges and pixels, by changing the input edges.
- 3- Decreasing the space between the pixel's parts.

This implies clearly the importance of the pre-processing on the accuracy (recognition rate). Feature extraction played important rule in canny edge detection to extract the features from images and remove the noise.

HMM's for classification the standard approach is:

- Separate the data into the data sets for each class.
- Train one HMM per class.
- On the test set compare the likelihood of each model to classify each window.

Supervised training: use statistics of known samples.

- Trainer knows underlying sequence for samples.
- Estimates transition/emission probabilities from samples database.

Unsupervised training (learning): update HMM parameters based on new samples.



## 5. Conclusions

In this paper, we introduced a new technique for image recognition. We proposed our technique based on intelligent features using efficient operator (Canny). The major contribution of the presented technique is the pre-processing stage. In the pre-processing, the input image's edge is fixed to four pixels and spacing between various parts of the pixels is fixed to a trial threshold. The pre-processing steps have enhanced the performance of the system by about 3 % as explained previously. Although pre-processing has simple and important steps but it simplifies the features and enhances the recognition rate. Our proposed technique showed the best performance compared with the other tested techniques in the sense of speed and recognition rate.

## Acknowledgements

We would like to express the deepest gratitude and sincere appreciation to Al-Azhar University and Canadian university for offering their help, support, and encouragement in accomplishing this paper.

## References

- [1] Nisha, Sandeep Dahiya. Face Detection and Expression Recognition using Neural Network Approaches. Global Journal of Computer Science and Technology: F Graphics & Vision Volume 15 Issue 3 Version 1.0, Year 2015.
- [2] Gurpreet Kaur, Manbir Sandhu and Purnima." Facial Recognition: Issues, Techniques and Applications"International Journal of Advanced Research in Computer Science and Software Engineering Volume 6, Issue 2, February 2016, pp. 508-512
- [3] [Navneet Jindal , Vikas Kumar, " Enhanced Face Recognition Algorithm using PCA with Artificial Neural Networks", Electronics and Communication Engineering Department Samalkha Group of Institutes, IJARCSSE, pp 864-872, 2013.
- [4] Manisha M. Kasar , Debnath Bhattacharyya and Taihoon Kim." Face Recognition Using Neural Network: A Review" International Journal of Security and Its Applications Vol. 10, No. 3 (2016), pp.81-100.
- [5] Suman kumar Bhattacharyya and Kumar rahul." Face Recognition by Linear Discriminant Analysis "International Journal of Communication Network Security, ISSN: 2231 - 1882, Volume-2, Issue-2, 2013.
- [6] P. Phaneendra, V. Ramachandran\and E. Srinivasa Reddy."Human Face Detection and Recognition using PCA" and DCT in HMM. International Journal of scientific Engineering and technology research ISSN 2319- 8885 Vol.04, Issue.35, Pages: 7080-7085, August-2015.
- [7] Shweta Mehta, Shailender Gupta, Bharat Bhushan." Face recognition using Neuro- Fuzzy Interface System "International Journal of Signal Processing, Image Processing and Pattern Recognition.Vol.7,No.1.July 2014 PP.339-344.
- [8] H. Chang, A. Koschan, B. Abidi, and M. Abidi. Fusing continuous spectral images for face

recognition under indoor and outdoor illuminants. Machine Vision and Applications, 21(2):201–215, February 2010.

- [9] Xiaoming Cheng and Wushan. Cheng. Facial Expression Recognition Based on Edge Detection. International Journal of Computer Science & Engineering Survey (IJCSSES) Vol.6, No.2, April 2015.
- [10] Tapas, David M. Mount, Nathan S. Netanyahu, Christine D. Piatko, and Ruth Silverman. 2002 "An Efficient k-Means Clustering Algorithm: Analysis and Implementation" IEEE transactions on pattern analysis and machine intelligence, vol. 24, no. 7.
- [11] Chien, J-T. & Liao, C-P. Maximum Confidence Hidden Markov Modeling for Face Recognition. Pattern Analysis and Machine Intelligence. IEEE Transactions Vol. 30, No 4, pp. 606-616, April 2008.
- [12] Mrs. Swati Y.Raut and Mrs.Dipti.A.Doshi. A Face Recognition System by Hidden Markov Model and Discriminating Set Approach. International Journal of Scientific & Engineering Research, Volume 6, Issue 1, January-2015.

## Biographies



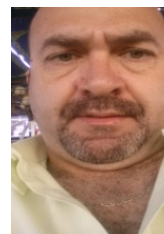
cairo.com , Tele: +20-01006551609

Sameh Abaza B.sc degree in Electronics and communication from 10th of Ramadan University 2005 .Finished master degree from Al-Azhar University 2009, Cairo, Egypt, in satellite image resolution enhancement technique and works as a lecturer in Electrical and communication department at Canadian international university. Email: sameh\_mohamed@cic-



signal processing, Digital Communication). Email: mohiyosof@yahoo.com.

Prof. Mohamed Ibrahim was born 1949. Finished Ph.D. jan1988 West Germany Ruhr University. He worked with Prof. Alfred, a famous Prof in Digital signal processing field .Worked Now as a Prof in Electrical and communication Engineer at Al-Azhar University, Egypt. He worked in Al-Azhar University from 1974. Published 174 research in different fields (Signal and image processing, Digital Filter and Algorithms for digital



International Symposium on signal processing and Information Technology, Email: Tarekmahmoud@hotmail.com

Tarek Mahmoud Ph.D. of electrical engineering, ECE department, University of Victoria, 2003, BC, Canada., M.Sc. of electrical engineering, Al-Azhar University, 1997, Cairo, Egypt. B.Sc. of electrical engineering. Ain- shams University, 1990, Cairo, Egypt. Distinction with First class honor. Published 4 researches in IEEE

